



Real-Time Rotation Estimation Using Histograms of Oriented Gradients

Blaž Bratanič^{1*}, Franjo Pernuš^{1,2}, Boštjan Likar^{1,2}, Dejan Tomažević^{1,2}

¹ Sensum, Computer Vision Systems, Ljubljana, Slovenia, ² Laboratory of Imaging Technologies, Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, Slovenia

Abstract

This paper focuses on real-time rotation estimation for model-based automated visual inspection. In the case of model-based inspection, spatial alignment is essential to distinguish visual defects from normal appearance variations. Defects are detected by comparing the inspected object with its spatially aligned ideal reference model. Rotation estimation is crucial for the inspection of rotationally symmetric objects where mechanical manipulation is unable to ensure the correct object rotation. We propose a novel method for in-plane rotation estimation. Rotation is estimated with an ensemble of nearest-neighbor estimators. Each estimator contains a spatially local representation of an object in a feature space for all rotation angles and is constructed with a semi-supervised self-training approach from a set of unlabeled training images. An individual representation in a feature space is obtained by calculating the Histograms of Oriented Gradients (HOG) over a spatially local region. Each estimator votes separately for the estimated angle; all votes are weighted and accumulated. The final estimation is the angle with the most votes. The method was evaluated on several datasets of pharmaceutical tablets varying in size, shape, and color. The results show that the proposed method is superior in robustness with comparable speed and accuracy to previously proposed methods for rotation estimation of pharmaceutical tablets. Furthermore, all evaluations were performed with the same set of parameters, which implies that the method requires minimal human intervention. Despite the evaluation focused on pharmaceutical tablets, we consider the method useful for any application that requires robust real-time in-plane rotation estimation.

Citation: Bratanič B, Pernuš F, Likar B, Tomažević D (2014) Real-Time Rotation Estimation Using Histograms of Oriented Gradients. PLoS ONE 9(3): e92137. doi:10.1371/journal.pone.0092137

Editor: Dewen Hu, College of Mechatronics and Automation, National University of Defense Technology, China

Received: December 18, 2013; **Accepted:** February 18, 2014; **Published:** March 24, 2014

Copyright: © 2014 Bratanič et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Ministry of Higher Education, Science, and Technology, Republic of Slovenia under grants L2-4072 and L2-5472, by Sensum, Computer Vision Systems, and by the European Union, European Social Fund. Boštjan Likar and Franjo Pernuš (two co-authors of this manuscript) are co-owners of the commercial funder of this research, Sensum, Computer Vision Systems. Boštjan Likar and Franjo Pernuš had a role in study design and preparation of the manuscript.

Competing Interests: The authors have the following interests: This work was partly supported by Sensum, Computer Vision Systems. Boštjan Likar and Franjo Pernuš are co-owners of Sensum, Computer Vision Systems while Blaž Bratanič and Dejan Tomažević are employed by Sensum, Computer Vision Systems. The funding agreement requires that the results of the research work are publicly available, however, Sensum can freely incorporate the results of the research in their commercial products. There are no patents, products in development or marketed products to declare. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials, as detailed online in the guide for authors.

* E-mail: blaz.bratanic@sensum.eu

Introduction

Quality control of an industrial process covers all aspects that influence the quality of a product [1]. It is based on measuring and assessing process parameters such as temperature and pressure as well as various product characteristics such as shape, hardness, composition, and visual appearance. Visual appearance can be important both for functional and aesthetic reasons and for compliance with statutory and regulatory requirements. Due to imperfect production processes, a compliant visual appearance is usually assured by a subsequent visual quality inspection [2–4].

Visual inspection is performed either manually by trained personnel or automatically by inspection machines. Manual inspection is slow, tedious, and subjective. It was reported [4] that human inspectors are prone to classify inspected objects as defective only to satisfy a rejection quota. Nevertheless, the high flexibility of a manual inspection still makes it a viable option for small production batches. Automatic inspection, by contrast, produces fast, objective, and reproducible results but requires a sophisticated system consisting of mechanical manipulation,

acquisition, registration, and analysis. Registration and analysis are difficult, because products can vary in size, shape, and surface complexity. Moreover, due to surface defects being visible only under directed illumination, registration and analysis must cope with rotation dependent object surface appearance.

Analysis only becomes tractable by the integration of *a priori* knowledge of the object, i.e., a reference model [3], [5]. Analysis is then performed by comparing the inspected object with its spatially aligned, ideal reference model. Rotation estimation is crucial for the inspection of rotationally symmetric objects where mechanical manipulation is unable to ensure correct object rotation.

We propose a method for model-based real-time rotation estimation. Rotation estimation uses an ensemble of nearest-neighbor estimators, where each estimator corresponds to a stationary spatially local region at various object rotations. Estimations from all estimators are weighted, based on the angle for which the estimator votes. Weighted estimations are then accumulated in a vote accumulator. The estimated angle is the angle with the maximum number of votes. Further, we describe

the estimator's weight estimation and propose a method to improve estimators from unlabeled training samples. The method was evaluated on several datasets of pharmaceutical tablets. At inspection, the tablets are mechanically constrained on a rotating drum, as shown in Figure 1. Rotation estimation of pharmaceutical tablets is a challenging task because of normal intra- and inter-tablet variability, which occurs due to imperfect production processes, the required generality of the method – it should work for a variety of shapes, imprints, and surface appearances – and speed requirements. A typical inspection machine inspects 100 tablets per second.

Although the evaluation is focused on pharmaceutical tablets, we consider the method useful for any application that requires robust real-time in-plane rotation estimation.

Related Work

Špiclin et al. [6] proposed three registration methods for in-plane rotation estimation: Direct Pixel Matching (DPM), Principal Axes Matching (PAM), and Circular Profile Matching (CPM). Direct Pixel Matching evaluates a similarity measure (SM) between the reference and a sample image. SM is a normalized cross-correlation of the rotating reference image and a sample image. The rotation angle ϕ^{est} is estimated by maximizing the SM between the two: $\phi^{est} = \arg \max_{\phi} SM(\phi)$. The method directly matches the pixels in $\{x, y\}$ space; thus the SM must be evaluated for each rotation angle individually. It therefore requires a rotation of the reference image. Each rotation includes an interpolation of the whole image, which induces a large performance penalty.

Circular Profile Matching, on the other hand, is based on the extraction and alignment of circular profiles. A circular profile is a 1-D function $I(\phi)$ obtained by integration of intensity values of a 2D image $I(x, y)$ within a ring centered at the tablet center:

$$I(\phi) = \int_{r_1}^{r_2} I(-r \sin \phi, r \cos \phi) dr. \quad (1)$$

The rotation angle is estimated by a 1-D cross correlation of the reference's and the sample's circular profiles. Compared to the



Figure 1. Mechanically constrained inspected products on an automated inspection machine.
doi:10.1371/journal.pone.0092137.g001

DPM, performance is improved by reducing the problem to one dimension.

Principal Axes Matching estimates the rotation between the reference and a sample image by matching the angles of the principal axes [7]. The principal axes are obtained on the assumption that an image $I(x, y)$ represents a density function $I: \mathbb{R}^2 \rightarrow \mathbb{R}$. Principal axes are the eigenvectors of the covariance matrix $C = E(I(x, y) \cdot (\mathbf{p} - \mathbf{m})(\mathbf{p} - \mathbf{m})^T)$, where m denotes the image center $[x_c, y_c]$ and $p = [x, y]$. The angle between the reference and sample image is obtained directly from the angles of the principal axes between the two images. Principal axes have a residual $\pm \pi$ sign ambiguity. To determine the rotation with principal axes reliably and accurately, matched images must also contain a distinctive shape asymmetry, and additional validation must be performed to resolve the $\pm \pi$ ambiguity. Špiclin et al. extensively evaluated the described methods and concluded that only the CPM method is suitable for rotation estimation of pharmaceutical tablets.

In general, prior work on rotation estimation is either sensitive to outliers due to non-robust global similarity measure [8], requires extensive parameter tuning, assumes rotationally symmetric illumination or does not run in real time.

Method

The key idea of the proposed method is to exploit the rotation dependence of HOG features to construct a representation of an object in a feature space for all rotation angles. Features are calculated on a dense grid of spatial regions, which tile a detection window. The detection window is determined in a preprocessing step from the boundaries of the segmented area [9]. Each spatial region corresponds to a separate reference set that contains a region appearance – in HOG feature space – at all rotation angles. Angle estimation is performed with an ensemble of nearest-neighbor (in Euclidean space) estimators, where each estimator individually votes for the resulting angle.

The method is divided into two phases: the training and the process phase. In the training phase, (Figure 2) reference sets describing object appearance on a valid rotation interval are constructed. A valid rotation interval is set by the user and determines valid object rotations (usually on the interval $[0, 2\pi)$). The rotation interval is discretized into a set of valid rotation angles Φ using a predefined discretization step δ_ϕ . Each reference set contains feature vectors describing the corresponding spatial regions at all valid rotation angles. Reference sets are initially constructed by rotating a single reference image and are then adjusted with a set of unlabeled training images. The unlabeled training set is also used to determine weights for each spatial region.

In the process phase, the constructed reference sets and estimated weights are used for angle estimation. Estimation is performed with an ensemble of nearest-neighbor estimators where each estimator votes for the resulting angle. All votes are weighted and accumulated. The resulting angle is the angle that accumulated most votes.

We begin with a brief description of Histograms of Oriented Gradients and then describe the construction of initial reference sets, their iterative adjustment, and a weight estimation scheme. We conclude with the description of a weighted voting-based rotation estimation.

Histograms of Oriented Gradients

Dalal and Triggs [10] proposed Histograms of Oriented Gradients as a feature set for robust human detection and

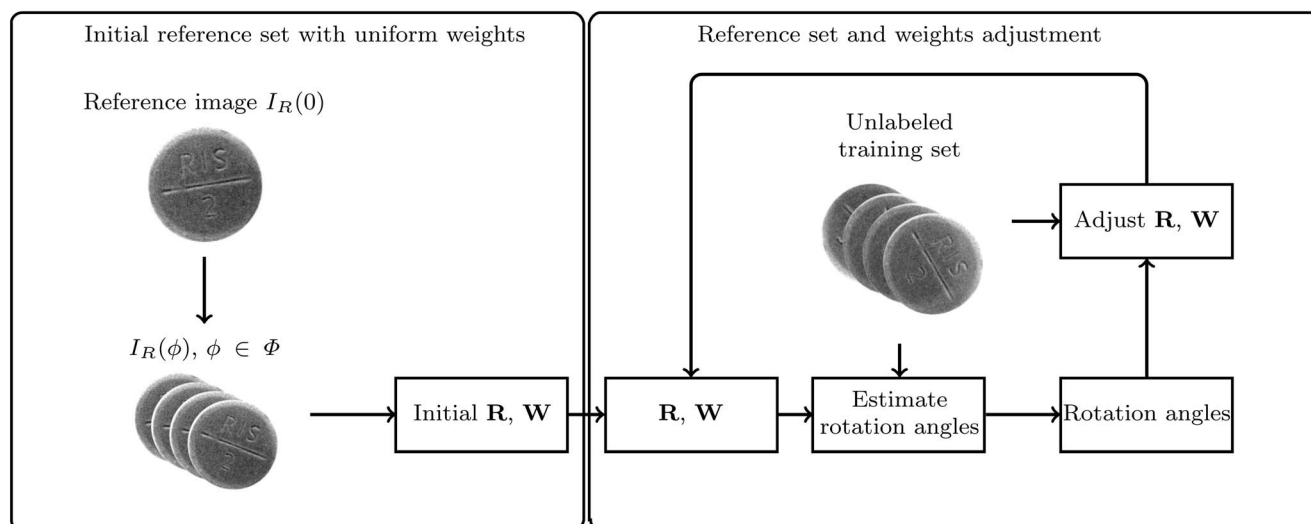


Figure 2. Training process. An initial weight and reference set estimations (left) are iteratively refined with a set of unlabeled training images (right). \mathbf{R} and \mathbf{W} denote reference and weights set respectively.
doi:10.1371/journal.pone.0092137.g002

localization. A feature set describes local appearance and shape by distributions of gradient orientations. The basic implementation defines an object boundary (detection window), sub-divided into smaller spatial regions (cells) (Figure 3). Each cell contains a distribution of gradient orientations. Illumination invariance is achieved by normalizing the cell values over larger blocks, which are formed by grouping cells in a sliding fashion. Zhu et al. [11] extended the original work to non-uniformly tiled cells with varying cell sizes and block layouts. Extracting features from overlapping cells by a straightforward implementation is time consuming because we redundantly evaluate the same areas multiple times. To avoid this, Zhu et al. [11] proposed an integral image approach to HOG feature extraction. Integral images (Figure 4) enable a constant time HOG calculation over an arbitrary rectangular region.

Reference set

We begin with the construction of a set of blocks \mathbf{B} over which we calculate HOG features. Each block consists of 4 cells in a 2×2 layout. To minimize user input, the blocks in a set \mathbf{B} span over multiple scales. On each scale, the blocks uniformly tile the detection window. We start at the largest scale – the whole detection window – and then iteratively reduce the size by a constant factor until we reach the smallest scale (Figure 5).

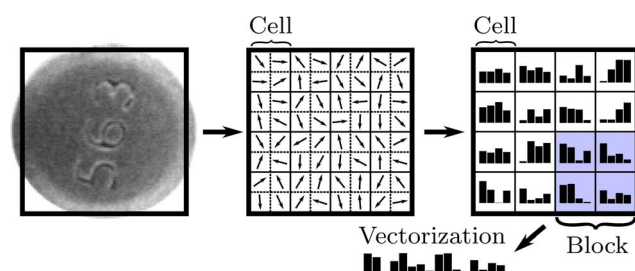


Figure 3. Histograms of Oriented Gradients. The detection window is subdivided into spatial cells, which are then concatenated into larger blocks.
doi:10.1371/journal.pone.0092137.g003

With a set of blocks $b \in \mathbf{B}$ and a reference image $I_R(\phi)$, a reference set \mathcal{R}^b is constructed for each block b .

Reference set $\mathcal{R}^b, b \in \mathbf{B}$, where $\mathcal{R}^b = \{\mathcal{R}_\phi^b\}, \phi \in \Phi$ is a set containing the features extracted from a block b over all valid angles Φ . A set of valid angles Φ is obtained with discretization of the valid rotation interval. The dimensionality of a feature vector \mathcal{R}_ϕ^b corresponds to the number of cells in each block and the number of histogram bins in each cell.

For all the angles $\phi \in \Phi$, we rotate the reference image by ϕ then calculate the HOG features over all blocks in \mathbf{B} . The resulting features are stored in corresponding reference sets \mathcal{R}^b .

Reference set adjustment. Initial reference sets, containing feature vectors describing an object appearance at all angles $\phi \in \Phi$, are constructed from a single image. This representation is biased, because it is constructed from a single reference image I_R assuming a rotationally invariant appearance. Bias is evident in a case of a corrupted reference image or in a case of rotationally non-symmetric illumination. The assumption of a constant surface appearance is relaxed by a semi-supervised [12] self-training approach. This enables the accommodation of an angle-dependent surface appearance, assuming only a piecewise constant surface appearance. By that we mean that on an arbitrarily small interval $[-\phi^{conf}, \phi^{conf}]$ the surface appearance remains constant.

The semi-supervised self-training approach is performed in an iterative fashion. Reference sets $\mathcal{R}^b, b \in \mathbf{B}$ are iteratively adjusted,

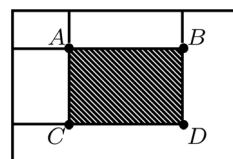


Figure 4. Integral image. The value $I(x, y)$ at any point (x, y) is the sum of all the pixels above and to the left of (x, y) inclusive, $I(x, y) = \sum_{x'=0}^x \sum_{y'=0}^y i(x', y')$ – integral F_{ABCD} of pixel values within arbitrary rectangle $ABCD$ is obtained by $F_{ABCD} = I(D) + I(A) - I(B) - I(C)$.
doi:10.1371/journal.pone.0092137.g004

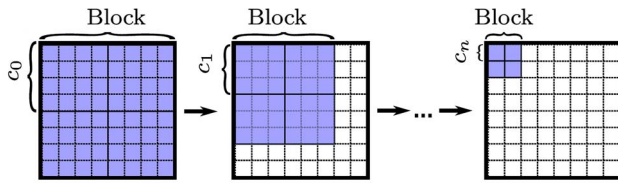


Figure 5. Initial set of cells. Cells are constructed by uniformly tiling the detection window with cells of decreasing size. Starting with whole detection window c_0 , we iteratively decrease the size by a constant factor $c_{n+1} = k c_n, 0 \leq k < 1$.
doi:10.1371/journal.pone.0092137.g005

with a set of n unlabeled training images $\mathcal{T} = \{T_i\}, i = 1, \dots, n$. Starting with the initial reference sets, we estimate angles ϕ_i^{est} for all training images T_i . For each training image T_i , we construct another set in HOG space $\mathcal{S}_i^b, b \in \mathcal{B}$, similar to the initial reference set \mathcal{R}^b , except that here each set describes the appearance only on the local interval $\phi \in [\phi_i^{est} - \phi^{conf}, \phi_i^{est} + \phi^{conf}]$.

This way we obtain multiple sets $\mathcal{S}_i^b, b \in \mathcal{B}, i = 1, \dots, n$, each containing a feature representation $\mathcal{S}_{i,\phi}^b$ of an object on a local interval (Figure 6).

Note that each element of a reference set $\mathcal{R}_\phi^b \in \mathcal{R}^b$ contains a feature representation of the corresponding block b at the specific angle $\phi \in \Phi$. Each element \mathcal{R}_ϕ^b is replaced with the median value of all object representations $\mathcal{S}_{i,\phi}^b, i = 1, \dots, n$, at the specific angle ϕ . The whole adjustment procedure is iteratively repeated until the angle estimations of the training images remain stationary or the maximum number of iterations is reached.

Weight estimation

In addition to the iterative reference set adjustment, we introduce an angle-dependent block-weighting scheme due to an empirical observation that not every block is equally relevant at all rotation angles (Figure 7). Weights quantify the importance of the local region $b \in \mathcal{B}$ when voting for the angle $\phi \in \Phi$.

Let \mathcal{W} be a set containing weights $\mathcal{W} = \{W_\phi^b\}, b \in \mathcal{B}, \phi \in \Phi$ for all the blocks \mathcal{B} at all valid angles Φ . The weight W_ϕ^b of block b and angle ϕ is estimated with an evaluation set \mathcal{E}^b , which is constructed by combining sets \mathcal{S}_i^b described in the previous section.

Each element in a set \mathcal{S}_i^b has a known angle ϕ_s on interval $[\phi_i^{est} - \phi^{conf}, \phi_i^{est} + \phi^{conf}]$ and represents an object appearance in

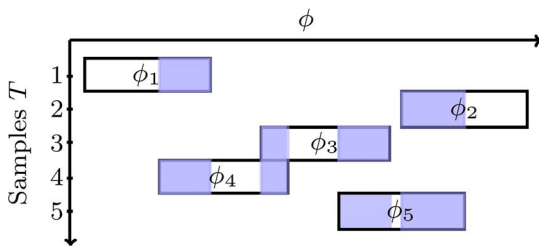


Figure 6. Reference set adjustment. Each of the training samples contributes to the reference set on local interval $\phi \in [\phi_i^{est} - \phi^{conf}, \phi_i^{est} + \phi^{conf}]$. For illustration purposes, angles with multiple representations are colored. Acquisition of unlabeled training images is usually trivial, thus an arbitrary number of feature representations can be obtained at every angle.
doi:10.1371/journal.pone.0092137.g006

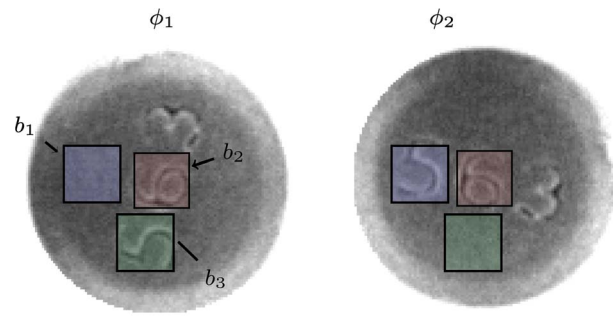


Figure 7. Set of blocks is equal for all rotation angles Φ . Each block contains relevant information, i.e., imprint, only at certain angles.
doi:10.1371/journal.pone.0092137.g007

the HOG feature space. The evaluation feature set for each block b is a combination of all local reference sets: $\mathcal{E}^b = \mathcal{S}_1^b \cup \dots \cup \mathcal{S}_n^b$.

For each element in a set \mathcal{E}^b with a known angle ϕ_s , k nearest neighbors are located in the reference set \mathcal{R}^b . Each nearest neighbor in the reference set has a corresponding angle $\phi_j^{ref}, j = 1, \dots, k$. The probability w_s^b that an element s in \mathcal{E}^b contributes to the true angle ϕ_s is a percentage of nearest neighbors in the α radius:

$$w_s^b = \frac{\sum_{j=1}^k 1[|\phi_s - \phi_j^{ref}| < \alpha]}{k} \quad (2)$$

The radius $\alpha = \frac{k\delta_\phi}{2}$ is calculated from the sampling resolution δ_ϕ of the reference set and a selected number of nearest neighbors k . Essentially, α is chosen such that at most k neighbors fit in the α radius.

A weight W_ϕ^b is obtained by collecting and averaging evaluation results w_s^b for all the elements in \mathcal{E}^b with the angle ϕ :

$$W_\phi^b = \frac{\sum_{s \in \mathcal{E}^b} w_s^b 1[\phi_s = \phi]}{\sum_{s \in \mathcal{E}^b} 1[\phi_s = \phi]} \quad (3)$$

This criterion should give higher weights to blocks that have a distance function with a well-defined global minimum. Figure 8, shows two distance functions. If the distance function $\text{dist}(\mathcal{R}^b, \mathcal{E}_s^b)$ — which is a Euclidean distance — has a well-defined global minimum (Figure 8a), most of the k -nearest neighbors will be in the α radius, resulting in high w_s^b . On the other hand, if a distance

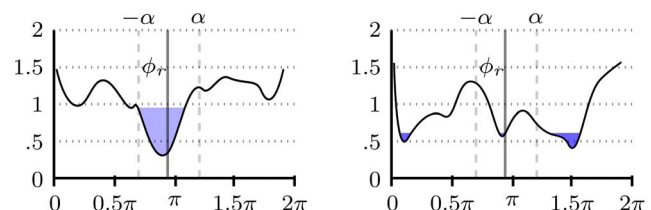


Figure 8. Nearest neighbors in an α radius. Left, all k nearest neighbors are in an α radius, by contrast most of the k nearest neighbors on the right figure are spread over the interval $[0, 2\pi)$.
doi:10.1371/journal.pone.0092137.g008

function does not have a well-defined global minimum (Figure 8b), k nearest neighbors will be spread over all the angles $\phi \in \Phi$, resulting in low w_s^b .

Rotation estimation

The rotation angle of the sample image I_s is estimated by approximate nearest-neighbor searches over the constructed reference sets \mathcal{R}^b . First, feature vectors over all the blocks $b \in B$ are extracted from the sample image I_s . Then, for each feature vector corresponding to the block $b \in B$, k nearest neighbors $\phi_j^b, j=1, \dots, k$ are located in the reference set \mathcal{R}^b . Each nearest neighbor contributes a single weighted-vote for its estimated angle.

All the votes are accumulated in an accumulator A_ϕ and weighted according to the weights in \mathcal{W} :

$$A_\phi = \sum_{b \in B} \sum_{j=1}^k W_\phi^b[\phi_j^b = \phi]. \quad (4)$$

The estimated sample's angle ϕ^{est} is then the angle with most accumulated votes:

$$\phi^{est} = \arg \max_{\phi} A_\phi. \quad (5)$$

Experiments and Results

Datasets

Eight datasets of various image quality and imprint visibility were obtained and captured for evaluation (Figure 9). Datasets were captured with a SPINE machine vision system (Figure 1), using a trilinear line-scan camera and fixed white LED illumination. The setup is shown in Figure 10. The eight datasets contain 514, 419, 272, 383, 447, 407, 329, and 329 images, respectively. Tablets differ in size, color, shape, and imprint.

Datasets 7 and 8 contain the same tablet type but were captured under a different illumination angle. Datasets 1–5 contain tablets with good imprint visibility; by contrast, datasets 6–8 contain images with poor imprint visibility, i.e., the angle is often hard to determine even upon close examination.

All datasets with supporting information are freely available upon request.



Figure 9. Evaluation datasets 1–8.
doi:10.1371/journal.pone.0092137.g009

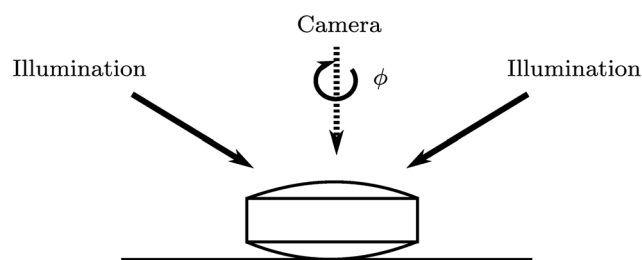


Figure 10. Illumination and camera view geometry. Camera view is parallel to the rotation axis.
doi:10.1371/journal.pone.0092137.g010

Implementation details

The proposed method along with the CPM method was implemented in C++. The HOG features were extracted using an integral image approach, due to blocks spanning over multiple scales of the detection window. An angle resolution δ_ϕ of the reference set was set to 1° . The same resolution was used for the similarity measure evaluation of the CPM. This effectively put the upper bound to estimation accuracy of both methods to 0.5° . A nearest-neighbor search in the reference set was performed using the FLANN [13] library and its kd-tree [14] approximate nearest-neighbor implementation. A separate kd-tree was constructed for each block, effectively speeding up the nearest-neighbor search by an order of magnitude, compared to an exhaustive search. For each block, $k=15$ nearest neighbors were located, weighted, and stored in the vote accumulator A_ϕ . The accumulator was smoothed with a $\sigma=1^\circ$ Gaussian kernel. The rotation angle was estimated with an exhaustive search over the vote accumulator. In the training phase, 160 unlabeled training images were used with a confidence interval $[-\phi^{conf}, \phi^{conf}]$, $\phi^{conf}=10^\circ$.

Metrics

The registration robustness and accuracy were evaluated by the criteria proposed by Špiclin et al. [6]. The registration robustness was determined by the percentage of successful registrations, i.e., registrations with an absolute error ϵ below 5° . Estimation accuracy is defined as the mean absolute angular error of successful registrations. Using this definition, inaccuracy has the upper bound at 5° . We found such restriction necessary to reduce the effect of outliers (results with large error) on the resulting accuracy. Otherwise, a single sample could have a disproportionately large effect on the resulting accuracy. Nevertheless, we quantified samples with large errors (samples with error ϵ above 5°) with hit percentage. A 5 degree upper bound was chosen to correspond with the upper bound proposed by Špiclin et al. [6].

An absolute error ϵ is a distance between the estimated angle ϕ^{est} and the gold standard angle ϕ^{ref} :

$$\epsilon = |\phi^{est} - \phi^{ref}| \quad (6)$$

The reference angle ϕ^{ref} is estimated from N manually positioned corresponding points on the reference $[x_l^r, y_l^r], l=1, \dots, N$ and sample $[x, y], l=1, \dots, N$ image. ϕ^{ref} is estimated by minimizing the mean squared distance, between the corresponding points as a function of rotation:

$$\phi^{ref} = \arctan \frac{\sum_{l=1}^N (x_l^s y_l^r - x_l^r y_l^s)}{\sum_{l=1}^N (x_l^r x_l^s + y_l^r y_l^s)} \quad (7)$$

Evaluation

The evaluation was performed with a one-round cross-validation. From each dataset, 160 unlabeled images were used for the training, and the rest of the images were used for the evaluation. First, we evaluated robustness, accuracy, and speed on eight captured datasets. Then, we evaluated the registration robustness of the proposed method on synthetically degraded samples. Synthetic degradation allowed for evaluation at a predefined degree of surface degradation. Finally, we evaluated the effect of the reference set adjustment and weighting of the nearest neighbor estimators (nn-estimators) on accuracy and hit percentage.

Evaluation on real datasets. Table 1 shows the registration accuracy and the hit percentage for the eight captured datasets. Average execution times for each method are shown in Table 2. Both the proposed method and the CPM method successfully estimated the angles, that is with error $< 5^\circ$, for all the samples in sets 3 and 5. For sets 1, 2, and 7, the CPM method failed on .8%, 1.4%, and 1.5% of samples, respectively. The proposed method, on the other hand, successfully estimated the angles for all the samples. Sets 4, 6, and 8 were more difficult datasets due to poor imprint visibility and $\pm \pi$ ambiguity. For set 4, the CPM method failed in 2.7% of cases and for set 8, in 1.5% of cases. The proposed method successfully estimated the angles for 100% of the samples. For set 6, a success rate of the CPM is 94%. The samples in this set had a poor imprint visibility (Figure 9) and a near symmetric imprint shape. For some samples, the rotation angle was hard to determine even upon close examination. Irrespective of this, the proposed method exhibited a near 100% success rate.

The accuracy of the proposed method was comparable on all evaluation datasets (Table 1). For sets 1, 2, 3, 6, 7 and 8, the proposed method achieved the highest accuracy; while for sets 4 and 5 the difference in accuracy was negligible ($\sim .3^\circ$). The accuracy of the proposed method was affected by accumulator smoothing. When smoothing the accumulator with the Gaussian kernel, we assumed that the votes were normally distributed with the mean at the true rotation angle. As shown in Figure 11, the vote distribution was skewed, thus smoothing introduced an estimation bias. Smoothing was, nevertheless, essential because of the limited number of votes.

Table 2. Average process time for compared methods.

	Process time [ms]
CPM	1.02
Proposed method	1.36

doi:10.1371/journal.pone.0092137.t002

Robustness evaluation. Two synthetically degraded datasets simulating defective surfaces were created by occluding a certain percentage of the tablet area. Various degrees of occlusion simulate various degrees of imprint defects. Datasets created in this way enable registration robustness and accuracy evaluation at a predefined degree of appearance changes.

The first set was generated by degrading samples in dataset 1 and the second set, by degrading samples in dataset 6. The tablet area on each sample was occluded at random positions with patches taken from the tablet area without imprint (Figure 12).

The evaluation of the first synthetically degraded dataset (Figure 13) showed that the success rate of the CPM declined to 90% at 10% of the sample area occluded. The proposed method achieved a 90% success rate even with 50% of the tablet area occluded. The evaluation of the second degraded dataset demonstrated similar results. Registration robustness is of utter importance when performing a model-based visual quality inspection, because a failed registration will result in wrong classification.

Effect of reference set adjustment and weighting of nn-estimators. The effect of the reference set adjustment and weighting of nn-estimators was evaluated on dataset 6; this set was used because it contained samples with high appearance variability and poor imprint contrast. Results are shown in Table 3. The results show that reference set adjustment is essential to increase registration robustness. The registration success rate without weighting and reference set adjustment was 87.4%. By introducing weights, the success rate increased to 94.9%; by adjusting the reference set, it increased to 99.3%. Combining the adjustment of the reference set and the weighting of the spatial regions, we obtained a 99.8% success rate. On this dataset, the CPM method achieved a 93% success rate.

Discussion

The main difficulty in the registration of pharmaceutical tablets arises from intra- and inter-tablet variability. Furthermore, a registration method should be general enough to work for a variety of shapes, imprints, and surface appearances, with minimal human intervention, while retaining sufficient accuracy, robustness, and

Table 1. Registration results for the proposed method and two reference registration methods.

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8
Error ($^\circ$)								
CPM	1.27	1.90	1.28	1.50	1.64	2.00	1.02	1.34
Proposed method	1.27	1.69	0.95	1.68	2.02	1.55	0.85	1.32
Hits (%)								
CPM	99.2	98.6	100.0	97.3	100	93.0	98.5	98.5
Proposed method	100.0	100.0	100.0	100.0	100	99.8	100	100.0

doi:10.1371/journal.pone.0092137.t001

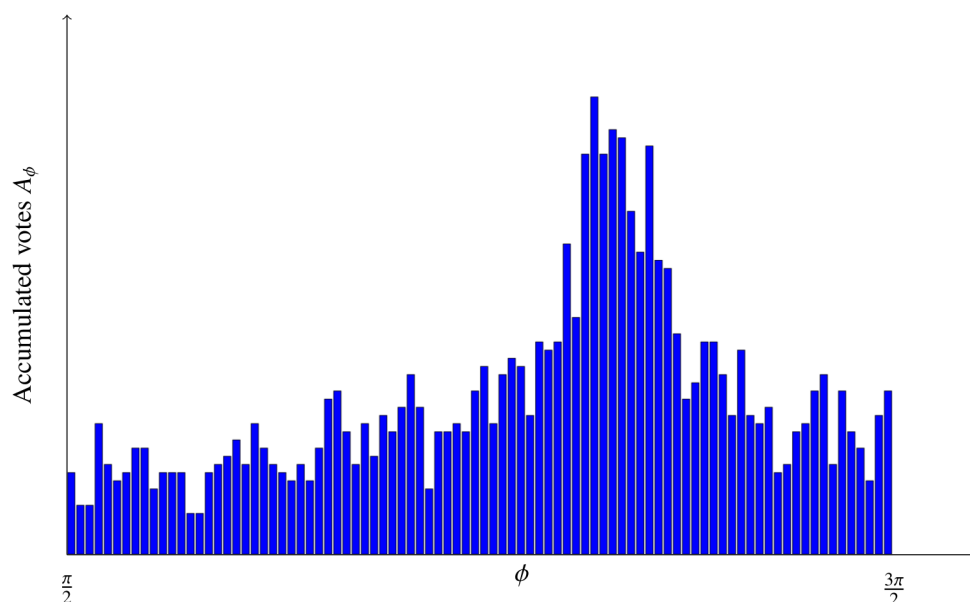


Figure 11. Distribution of accumulated votes is skewed. For illustration purposes only an interval $[\frac{\pi}{2}, \frac{3\pi}{4}]$ is shown.
doi:10.1371/journal.pone.0092137.g011

performance for real-time inspection (< 5 ms). A method based on the extraction of circular profiles [6] had been proposed previously. Its main advantages were simplicity and speed, given the assumption of rotationally symmetric illumination and rotationally invariant object appearance. It required a parameter fine-tuning for each tablet type individually, an accurate segmentation, and an accurate center estimation (estimation error below 1% of an object size).

In practice, perfect diffuse illumination is undesirable for the inspection of surface structures, because shapes and surface variations are difficult to recognize. Under directed illumination, on the other hand, surface appearance varies with the angle of rotation.

In the proposed method – without reference set adjustment – a single reference image is used to build a reference set containing an object representation for all valid rotation angles. However, under directed illumination, a single reference image is not a good

description of the surface appearance at all rotation angles, because appearance varies with the angle of rotation. The proposed method automatically adjusts a representative model of the object, avoiding the rotationally invariant appearance requirement. This is done by a semi-supervised self-training approach, iteratively improving the object representation at various angles. Due to a non-parametric approach to angle estimation this is easily incorporated into the estimation. The method builds the object representation assuming a locally constant appearance on an arbitrarily small interval $[-\phi^{conf}, \phi^{conf}]$. As already noted in [6], in-plane rotations of round tablets are random and uniformly distributed on $[0, 2\pi)$

interval. Given the uniform distribution, we obtain $\frac{n\phi^{conf}}{\pi}$ different feature representations at each angle. Here n is the number of training samples. Clearly, an arbitrary number of feature representations can be obtained for an arbitrarily small ϕ^{conf} by increasing the number of training samples. This is a feasible approach, because the acquisition of a large set of unlabeled samples is usually simple to do and economical. An improved representation of an object is the main reason for the improvement of registration robustness.

Registration error increases slightly with reference set adjustment. This is because we use a finite number of unlabeled training samples for the reference set adjustment. To incorporate information of each unlabeled sample, a rotation angle for each sample must first be estimated using the unadjusted reference set. Consequently, we are adjusting the reference set with training samples with the estimated rotation angles, where some estimations can be erroneous. Erroneous estimations affect the adjustment process and subsequently introduce error into the estimation framework.

By decoupling spatial regions and by treating each as an individual model, the dimensionality of each model is reduced, consequently enabling the use of spatial structures for a fast nearest-neighbor search. Decoupling also increases estimation robustness by eliminating the effects of outliers but reduces the discriminant power of each model. This approach closely



Figure 12. Synthetically degraded dataset. Occlusion patches are randomly spread over the tablet surface. Figures show samples with 0–50% of the area occluded.
doi:10.1371/journal.pone.0092137.g012

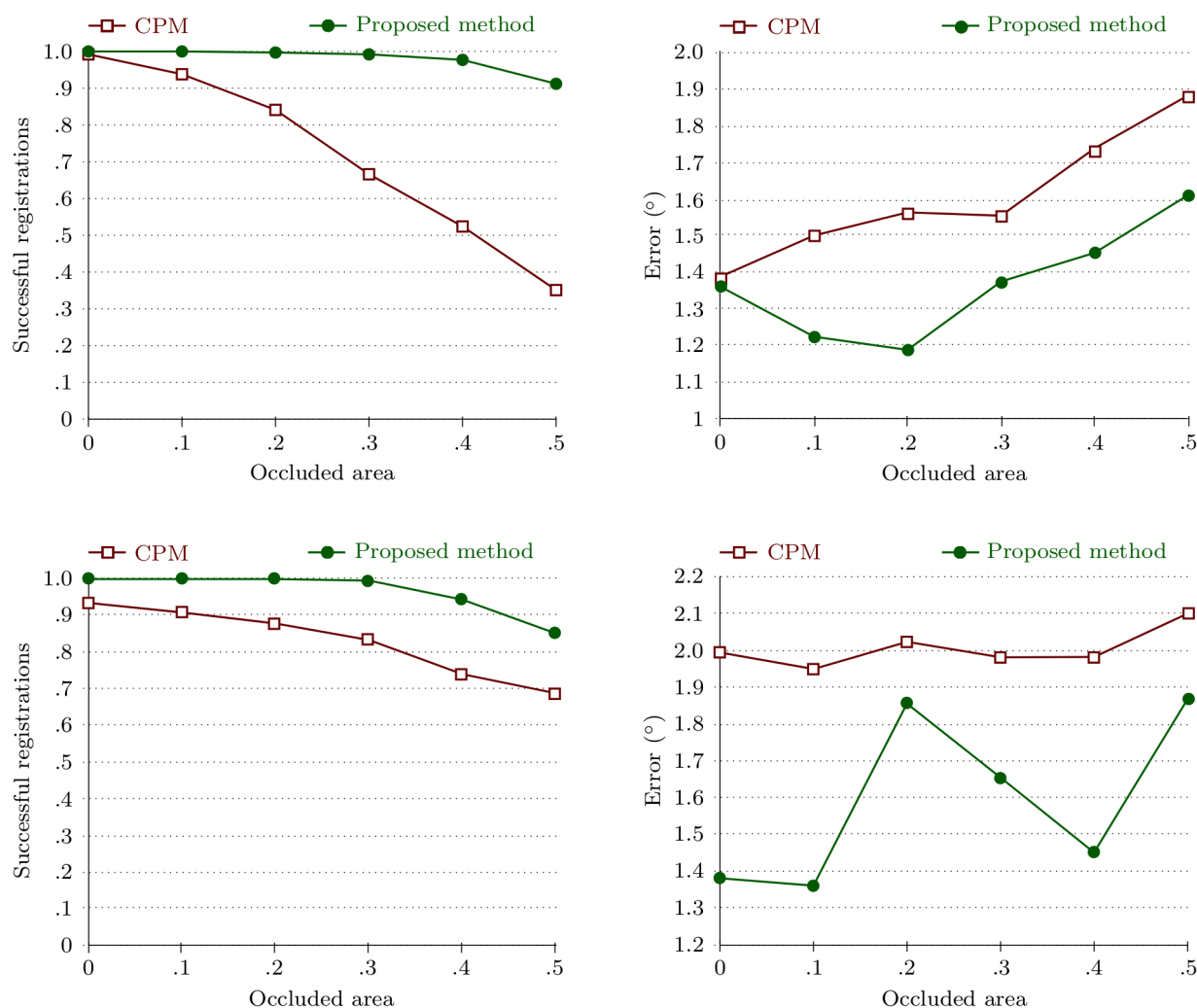


Figure 13. Robustness (left column) and accuracy (right column) evaluation, performed on synthetically degraded images. Results for first and second synthetically degraded datasets are presented in the first and second row respectively.
doi:10.1371/journal.pone.0092137.g013

resembles the description of the ensemble methods given in [15]. The advantage of the ensemble approach is shown by increased robustness and accuracy [16]. In the proposed method, each model determines an angle solely from its local neighborhood. This increases the robustness but reduces the discriminative power of each model. This reduction is alleviated by adjusting the

reference set to represent better an object at all rotation angles, and by weighting each spatial region according to the angle for which it votes for. Figure 14 illustrates the weight of each spatial region at various angles for two different tablet types. Essentially, both weighting and reference set adjustment achieve a similar goal by different means. Both increase the robustness of the method by reducing the power of irrelevant spatial regions. To highlight the advantage of the proposed method further, note that all the results are obtained with the same set of parameters, whereas the CPM method requires parameter fine-tuning for each dataset individually.

In the proposed method, we used a uniform grid partitioning to calculate the HOG feature descriptors. Nonetheless, other partitioning strategies are also possible. For natural images, Zhou et al. [17] partitioned an image into several rectangular horizontal and vertical partitions. On natural images, many objects remain in vertically and horizontally consistent locations relative to other objects; note that this does not necessarily hold true for other types of images. In [11], Zhu et al. used the Adaboost algorithm to select an optimal set of blocks for human detection. In our work, we did not consider this approach, because boosting algorithms [18],

Table 3. Effect of block weighting and reference set adjustment on registration accuracy.

	Uniform weights	Non-uniform weights
Error (°)		
Initial	1.52	1.51
Adjusted	1.59	1.55
Hits (%)		
Initial	87.4	94.9
Adjusted	99.3	99.8

doi:10.1371/journal.pone.0092137.t003

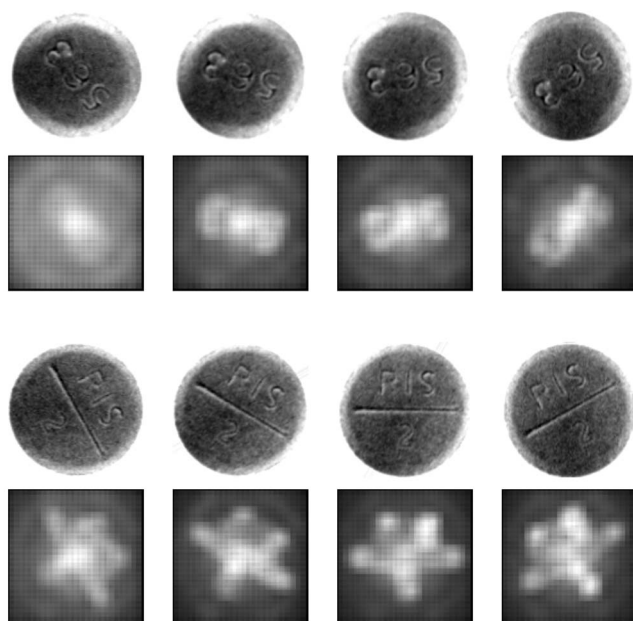


Figure 14. Cell weight visualization for different rotation angles. Higher intensity values represent higher weights.
doi:10.1371/journal.pone.0092137.g014

while common for feature selection, usually require extensive training times and large training sets.

In the context of the current state-of-the-art methods for rigid-body registration, the proposed method uses a dense grid for feature calculation, with a finite set of possible transformations. This approach imposes an upper bound to registration accuracy. By contrast, most state-of-the-art methods start by detecting local regions invariant under affinity, and then calculate features over those regions. Several region detectors were proposed for this task

[19], mostly focusing on region detection on planar structures. A descriptor [20] is associated with each region, which is used to establish correspondences between the regions and their descriptors on multiple images. Region detection and correspondence search are difficult in the presence of repetitive or uninformative structures. With a dense approach, both region detection and correspondence search are avoided, thus eliminating two possible points of failure. This is possible because we are concerned only with a single degree of freedom – not the whole affine transformation. Note, that in medical imaging numerous robust methods [21] have been proposed for rigid [22] [23] and non-rigid [24] registration. However, the application of those methods in industrial setting is usually impractical due to high computational complexity.

Conclusion

In this paper, we proposed a new rotation estimation method and demonstrated its performance and efficiency relative to another method for rotation estimation of pharmaceutical tablets. The proposed method requires minimal human intervention; all evaluation results were obtained with the same set of parameters. In addition, it does not require segmentation, because the voting scheme reduces the voting power of outliers, e.g., wrongly segmented or defective regions. Using a real dataset, we show that introducing unlabeled training images increases the robustness by relaxing the constant appearance assumption. Despite the fact that the evaluation focused on pharmaceutical tablets, we consider the method useful for any application that requires robust real-time rotation estimation.

Author Contributions

Conceived and designed the experiments: BB DT BL FP. Performed the experiments: BB. Analyzed the data: BB DT. Wrote the paper: BB DT BL FP. Designed the software used in analysis: BB.

References

- Bunn G (2006) Good manufacturing practices for pharmaceuticals, volume 169. Informa Healthcare.
- Chin RT, Harlow CA (1982) Automated visual inspection: A survey. *IEEE Trans Pattern Anal Mach Intell* vol. 4: 557–573.
- Bukovec M, Žiga Špiclin, Pernuš F, Likar B (2007) Automated visual inspection of imprinted pharmaceutical tablets. *Measurement Science and Technology* 18: 2921.
- Newman TS, Jain AK (1995) A survey of automated visual inspection. *Computer Vision and Image Understanding* 61: 231–262.
- Možina M, Tomažević D, Pernuš F, Likar B (2011) Automated visual inspection of imprint quality of pharmaceutical tablets. *Mach Vision Appl* 1–11.
- Špiclin Ž, Bukovec M, Pernuš F, Likar B (2011) Image registration for visual inspection of imprinted pharmaceutical tablets. *Mach Vision Appl* 22: 197–206.
- Alpert NM, Bradshaw JF, Kennedy D, Correia JA (1990) The principal axes transformation—a method for image registration. *J Nucl Med* 31: 1717–1722.
- Goshtasby AA (2012) *Image Registration: Principles, Tools and Methods*. Springer-Verlag New York Incorporated.
- Možina M, Tomažević D, Pernuš F, Likar B (2011) Real-time image segmentation for visual inspection of pharmaceutical tablets. *Mach Vision Appl* 22: 145–156.
- Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*. volume vol. 1, pp. 886–893.
- Zhu Q, Yeh MC, Cheng KT, Avidan S (2006) Fast human detection using a cascade of histograms of oriented gradients. In: *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*. volume vol. 2, pp. 1491–1498.
- Zhu X (2005) Semi-supervised learning literature survey. Technical Report 1530, University of Wisconsin, Madison.
- Muja M, Lowe DG (2009) Fast approximate nearest neighbors with automatic algorithm configuration. In: *International Conference on Computer Vision Theory and Application VISSAPP'09*. INSTICC Press, pp. 331–340.
- Bentley JL (1975) Multidimensional binary search trees used for associative searching. *Communications of the ACM* 18: 509–517.
- Mendes-Moreira J, Soares C, Jorge AM, Sousa JFD (2012) Ensemble approaches for regression: A survey. *ACM Computing Surveys (CSUR)* 45: 10.
- Garcia-Pedrajas N, Hervás-Martínez C, Ortiz-Boyer D (2005) Cooperative coevolution of artificial neural network ensembles for pattern classification. *Evolutionary Computation, IEEE Transactions on* 9: 271–302.
- Zhou L, Zhou Z, Hu D (2013) Scene classification using a multi-resolution bag-of-features model. *Pattern Recognition* 46: 424–433.
- Ferreira AJ, Figueiredo MA (2012) Boosting algorithms: A review of methods, theory, and applications. In: *Ensemble Machine Learning*, Springer. pp. 35–85.
- Mikolajczyk K, Tuytelaars T, Schmid C, Zisserman A, Matas J, et al. (2005) A comparison of affine region detectors. *Int J Comput Vision* 65: 43–72.
- Mikolajczyk K, Schmid C (2005) A performance evaluation of local descriptors. *IEEE Trans Pattern Anal Mach Intell* 27: 1615–1630.
- Maintz J, Viergever MA (1998) A survey of medical image registration. *Medical image analysis* 2: 1–36.
- Hill DL, Batchelor PG, Holden M, Hawkes DJ (2001) Medical image registration. *Physics in medicine and biology* 46: R1.
- Chen F, Su L, Liu Y, Hu D (2013) Confirming the diversity of the brain after normalization: An approach based on identity authentication. *PLoS one* 8: e54328.
- Holden M (2008) A review of geometric transformations for nonrigid body registration. *Medical Imaging, IEEE Transactions on* 27: 111–128.